

ČESKÉ NORMATIVNÍ SKÓRY BECKOVY ŠKÁLY DEPRESIVITY: METAANALÝZA

JAROSLAV GOTTFRIED¹

¹ Katedra psychologie, Fakulta sociálních studií Masarykovy univerzity

Abstrakt: Úvodem je představena historie lokální adaptace a standardizace Beckovy škály depresivity spolu se sociodemografickými proměnnými potenciálně ovlivňujícími získané skóry. Jádrem studie se skládá z rešerše dosavadních výzkumů předkládajících českou adaptaci metody neklinickým či klinicky depresivním lidem. Získané deskriptivní statistiky byly použity pro metaanalytický odhad populačních parametrů skóre. Výsledky metaanalýzy ukazují vliv formy administrace na výsledný skóre, kdy online forma způsobuje průměrné zvýšení skóre o více než tři body. I přes kontrolu formy administrace lze napříč studiemi sledovat výraznou heterogenitu průměrných skóre a směrodatných odchylek, která komplikuje odhad populačních parametrů. Pravděpodobnými příčinami rozdílů je výběrové zkreslení vzorku a kontextové efekty administrace. Kvůli výrazné heterogenitě skóre výzkumných vzorků je do budoucna žádoucí prozkoumat možný vliv těchto zkreslení i u jiných metod. Na základě odhadnutých populačních parametrů a teoretických předpokladů bylo následně simulováno rozložení populačních dat, které bylo použito pro stanovení přibližných cut-off skóre pro diagnostické účely. Na závěr jsou výsledky porovnány s výsledky zahraničních metaanalýz a jsou shrnuta doporučení ohledně používání Beckovy škály depresivity pro budoucí výzkum i pro praktickou diagnostiku. Závěry studie limituje především malé množství studií zahrnutých do metaanalýzy, heterogenita skóre napříč studiemi a nesplněný předpoklad normálního rozložení skóre.

Klíčová slova: Beckova škála; BDI-II; deprese; metaanalýza; normy

Teoretický úvod

Beckova škála depresivity (zkr. BDI-II; Beck, Steer, & Brown, 1996) je 21položkový psychodiagnostický nástroj sloužící ke screeningovému vyšetření aktuální závažnosti depresivních symptomů u dospělých, obvykle metodou tužka-papír. Česká adaptace BDI-II (Preiss & Vacíř, 1999) dlouho postrádala důslednou standardizaci a ověření psychometrických vlastností (viz Gottfried, 2015), nicméně v horizontu let 2016–2019 se situace v tomto ohledu významně zlepšila. V uvedeném období byly publikovány čtyři

¹ Katedra psychologie, Fakulta sociálních studií MU, Joštova 10, 602 00, Brno
Doručeno do redakce: 23. 6. 2019

studie, které lze označit za standardizační či normativní. Jedná se o články Ptáčka, Rabocha, Vňukové, Hlinky a Anderse (2016), Ociskové et al. (2017) a diplomové práce Čihařové (2017) a Pánka (2018).

Nové studie přinášejí tolik potřebné informace o psychometrických kvalitách české adaptace BDI-II, jako jsou normativní data s cut-off skóry a důkazy o validitě a reliabilitě, které až donedávna chyběly, díky čemuž poskytují podle mého názoru dostatečný podklad pro užívání BDI-II v praxi. Na obecné rovině studie dospívají k podobným, pozitivně laděným závěrům ohledně kvality BDI-II coby metody pro screeningové vyšetření depresivních symptomů, nicméně reportované rozložení skóre v populaci a hypotetizované faktorové struktury nástroje se napříč výše čtyřmi citovanými studiemi liší. Rovněž kvůli faktu, že diplomové práce Čihařové (2017) a Pánka (2018) nediskutují výsledky předchozích standardizačních studií, působí nové výzkumy BDI-II, ač nepochybně záslužné a přínosné, roztržtým dojmem, jako by se odehrávaly ve svém vlastním prostoru, odděleně od ostatních. V současnosti tak chybí jakákoliv výraznější integrace informací z dosavadních tuzemských studií české adaptace BDI-II. Systematickou integrací informací lze kromě komplexnějšího poznání dospět také k lepším doporučením pro praktickou psychodiagnostiku. V neposlední řadě zmiňovaná integrace usnadňuje čtenářům z řad výzkumníků i praktiků orientaci v tématu v porovnání s alternativou, kdy si čtenář musí studie i jejich výsledky dohledávat, srovnávat a případné nesrovnalosti podle svých vlastních možností interpretovat.

Cílem předkládané studie je poskytnout právě takovou integraci dosavadních informací ohledně metody BDI-II, konkrétně informací o průměrných skórech neklinické a klinické populace a jejich variabilitě v jednotlivých studiích, stejně jako doplňujících informací o sběru dat a charakteristikách výzkumného vzorku, díky čemuž bude možné odhadnout rozložení skóre v obecné populaci a případně i faktory, které jej ovlivňují. Přestože se česká adaptace BDI-II již neprodává, je v různých podobách snadno dostupná na internetu a je stále hojně používána ve výzkumu i v praxi (Urbánek, 2010). Z tohoto důvodu vidím jako užitečné se této metodě výzkumně věnovat a provést metaanalýzu informací získaných nejen ze standardizačních studií BDI-II, nýbrž i ze studií a diplomových prací, které BDI-II využily pro různorodé účely, například jako měřítko kriteriální či souběžné validity při standardizaci jiné metody.

V metaanalytických studiích je možné a mnohdy vhodné kontrolovat vliv intervenujících proměnných. V souvislosti s metodou BDI-II se nabízí kontrolovat vliv celkem tří demografických proměnných: věku, pohlaví a vzdělání. Přestože vyšší skóry BDI-II u starších participantů dávají teoretický smysl, jelikož položky BDI-II cílí i na somatické faktory, například chuť k jídlu či cirkadiánní rytmus, studie od Vrané (2012), Ociskové et al. (2017), Čihařové (2017), ani Pánka (2018) nenalezly vliv věku na skóre BDI-II a Preiss, Vraná a Kuchařová (2010) našli pouze slabou pozitivní korelaci ($r = 0,38$), která se však při současné kontrole pohlaví v regresní analýze neukázala být statisticky signifikantní. Věk participantů pro potřeby této metaanalýzy proto neberu v potaz.

Ani vliv pohlaví nekontroluji. Důvodem jsou smíšené výsledky předchozích studií s obecně malou silou účinku. Například Whisman, Judd, Whiteford a Gelhorn (2013) uvádějí, že původní anglická verze BDI-II měří muže a ženy srovnatelně, minimálně co se týče faktorové struktury metody. Z tuzemských výzkumů pak Preiss et al. (2010c), Ocisková et al. (2017) a Pánek (2018) rozdíl mezi pohlavími nenalézají, avšak Vraná (2012) i Ptáček et al. (2016) ano, konkrétně zhruba o dva body vyšší součtové skóry BDI-II u žen, a stejně tak Čihařová (2017) ve skupině žen pozoruje přibližně o jeden bod vyšší skóry. Smíšené výsledky poskytují též zahraniční studie (viz Čihařová, 2017, s. 54, Tabulka 3). Z uvedených výzkumů lze soudit, že skór české adaptace BDI-II s pohlavím spíše nesouvisí a pokud ano, tak pouze slabě a lze očekávat, že ženy budou průměrně dosahovat skórů vyšších o maximálně dva body oproti mužům.

Co se týče vlivu vzdělání, Čihařová (2017) pozorovala, že lidé s vysokoškolským vzděláním průměrně skórovali o přibližně pět bodů méně než ti se středoškolským vzděláním, a Pánek (2018) uvádí rozdíl mezi týmiž skupinami necelé čtyři body. Rovněž Preiss et al. (2010c) našli statisticky signifikantní spojení vyššího vzdělání a nižšího skóru depresivity, bohužel bez uvedené velikosti účinku. Ačkoliv Ocisková et al. (2017) nenalezli signifikantní rozdíly ve skórech BDI-II podle vzdělání, z jejich deskriptiv je zřejmý trend klesajícího skóru lidí s vyšším vzděláním, avšak pouze v kontrolní skupině, kde lze pozorovat výrazný rozdíl 12 bodů mezi průměrnými skóry lidí se základním a vysokoškolským vzděláním. Nález nižších skórů depresivity u lidí s vyšším vzděláním vybízí k několika možným interpretacím. Lidé s vyšším vzděláním mohou k vyplňování BDI-II přistupovat odlišně, například jinak chápat význam některých položek, uvažovat nad pojmy v jiném kontextu nebo volit specifický odpověďový styl. Nižší skóry mohou být i důsledkem odlišných životních podmínek lidí s různou úrovní vzdělání, odrážejících se v jejich socioekonomickém statusu, přístupu ke zdravotní péči i péči o duševní zdraví, nebo zdrojích sociální opory. Rovněž je možné, že lidé s vyšším vzděláním mají lepší kognitivní schopnosti a sociálně-emoční dovednosti, díky čemuž uplatňují obecně lepší copingové strategie a jejich řešení problémů je efektivnější, což může fungovat jako účinná prevence stresových a depresivně laděných stavů. Předložené nápady mohou posloužit jako odrazový můstek pro budoucí studie zkoumající příčiny a implikace spojení vyššího vzdělání s nižší depresivitou, nicméně toto téma není v ohnisku zájmu této studie, a proto se jím zde podrobněji nezabývám.

I když by bylo vhodné vliv vzdělání do metaanalýzy zahrnout, po pilotní rešerši literatury jsem zjistil, že s výjimkou klinických studií většina z celkového počtu relevantních výzkumů nekontroluje vzdělání probandů, což je pochopitelné, jelikož vzdělání často nebylo v centru výzkumného zájmu. Při absenci údajů vliv vzdělání napříč studiemi nemohu kontrolovat, a proto informace o vzdělání participantů do metaanalýzy nezahrnuji kvůli jejich nedostupnosti. Z výše uvedených důvodů považuji v rámci metaanalýzy vliv věku, pohlaví a vzdělání na součtový skór BDI-II nekontroluji, ačkoliv jej především v případě vzdělání očekávám.

Zamýšleným výstupem mé studie má být vytvoření robustních norem české adaptace BDI-II pro zdravou a depresivní populaci na základě dostupných informací o průměru a směrodatných odchylkách skóre vzorků z dosavadních výzkumů používajících tuto metodu. Jinými slovy je cílem studie odhadnutí průměrů a směrodatných odchylek populačního rozložení zdravé a depresivní populace. Skrze metaanalýzu rovněž plánuji zhodnotit variabilitu skóre mezi jednotlivými studiemi a na základě její velikosti vyvodit důsledky pro výzkum i pro praxi.

Metoda

Rešerše metod

Pro analýzu klíčových parametrů bylo mým záměrem shromáždit data o průměrných součtových skórech české adaptace BDI-II a jejich směrodatných odchylkách na neklinické a klinické populaci. Při rešerši literatury jsem primárně využil rozhraní Google Scholar s klíčovými frázemi: BDI-II, standardizace, Beckova škála, deprese. Sekundárně jsem zdroje manuálně dohledal skrze seznamy literatury v relevantních výzkumech a skrze funkci Google Scholar „Počet citací tohoto článku“ zobrazující, kterými studiemi je vybraný výzkum citován. Relevantní studie byly vybírány na základě splnění následujících podmínek:

1. Ve studii je použita česká adaptace metody BDI-II podle Preisse a Vacíře (1999) bez výrazných dodatečných úprav
2. Studie nezahrnuje experimentální manipulaci či kontext vyšetření s potenciálem ovlivnit skóry BDI-II
3. Vzorek studie sestává převážně z českých či slovenských probandů a zároveň buď:
 - a) Může být považován za vzorek populace bez klinické diagnózy
 - b) Může být považován za vzorek klinické populace s depresivním onemocněním z kategorie F32 nebo F33

Pokud studie splňovala všechna kritéria, vyhledal jsem v ní informace o počtu participantů, průměrném skóru BDI-II a jeho směrodatné odchylce. V případě, že údaje nebyly ve studii dohledatelné, požádal jsem skrze e-mail kontaktního autora o přístup k primárním datům, či alespoň o poskytnutí chybějících údajů. Studie, jejichž klíčové údaje jsem nebyl schopen získat, jsem z konečného seznamu analyzovaných zdrojů vyřadil. V případech, kdy studie pouze reportovala údaje pro více skupin respondentů zvlášť, bylo potřeba spojit tyto informace a vytvořit průměr a směrodatnou odchylku celého vzorku. Vzorec 1 vyjadřuje postup výpočtu celkového průměru, zatímco Vzorec 2 vyjadřuje postup výpočtu celkové směrodatné odchylky. Pro samotný výpočet jsem využil online aplikaci².

² Volně dostupnou na internetové adrese

https://www.statstodo.com/CombineMeansSDs_Pgm.php

Vzorec 1 – výpočet celkového průměrného skóru

$$M_T = \frac{\sum m_i n_i}{\sum n_i}$$

Vzorec 2 – výpočet celkové směrodatné odchylky

$$SD_T = \sqrt{\frac{\sum (sd_i^2 (n_i - 1) + (m_i n_i)^2 / n_i) - \sum (m_i n_i)^2 / \sum n_i}{\sum n_i - 1}}$$

V Tabulce 1 představuji výčet základních parametrů 22 vybraných studií s BDI-II skóry neklinických vzorků a v Tabulce 2 pak skóry 11 studií na klinických vzorcích s převážně unipolární depresí.

Tabulka 1 – studie reportující skóry BDI-II neklinických vzorků

Studie	Povaha vzorku	Věk	N	M	SD
Benda a Reichová (2016)	příležitostný online	28±10,2	415	11,8	9,41
Bob et al. (2005)	příležitostný	39,2±12,4	50	1,9	3,10
Bob et al. (2009)	příležitostný	34,1±7,1	35	5,3	5,40
Bob et al. (2010)	příležitostný	33,6±8,4	86	7,0	5,80
Bob et al. (2015)	data projektu ELSPEC	18-19	364	7,9	8,02
Čihařová (2017)	kvótní, příležitostný	18-96	450	7,0	5,80
Daňšová et al. (2016)	příležitostný online	24,7±6,4	1654	11,4	9,43
Divíšková (2015)	nenáhodný, náhodný, snow-ball	45,5±10,6	100	7,1	7,36
Dočkalová (2017)	nenáhodný, příležitostný	17-20	396	15,5	9,78
Janečková et al. (2013)	nenáhodný online, VŠ studenti	21,7±1,7	140	9,1	7,12
Nikolla (2018)	nenáhodný	33±15,8	127	7,0	6,29
Ocisková et al. (2016)	snow-ball	27,1±11,7	394	7,9	7,00
Ocisková et al. (2017)	snow-ball	27±10,26	767	7,2	6,80
Pánek (2018)	snow-ball, příležitostný	30,2±15,2	74	11,6	11,00
Pánek (2018)	snow-ball, příležitostný online	24,4±4,52	232	15,7	12,04
Preiss et al. (2009)	nenáhodný (vyvážení klin. skupiny)	46,1±12,8	97	6,6	5,97
Preiss et al. (2010a)*	nenáhodný (vyvážení klin. skupiny)	44,5±13,5	31	4,5	4,60
Preiss et al. (2010b)	nenáhodný (vyvážení klin. skupiny)	46,2±12	89	6,3	5,70
Preiss et al. (2010c)	kvótní, příležitostný, snow-ball	40,4±17,7	200	10,6	7,80
Preiss a Vacíř (1999)	příležitostný	18	60	10,3	7,02
Ptáček et al. (2016)	kvótní online (STEMMARK)	25-50	675	11,5	9,91
Ptáček et al. (2016)	kvótní face-to-face (STEMMARK)	51-65	352	9,4	9,36
Světlák et al. (2010)	příležitostný, VŠ studenti	22,1±1,9	196	7,7	5,60
Štěpánková Georgi et al. (2018)	příležitostný	67,9±6,6	86	9,7	6,27
Vraná (2012)	data výzkumného záměru MZČR0PCP2005	45,4±16,4	387	8,4	6,92

* studie uvádí chybné deskriptivní údaje, hodnoty v tabulce jsou odvozeny z originálního datasetu autorů

Z obou tabulek je již na první pohled patrné, že průměrné součtové skóry BDI-II v klinických i neklinických vzorcích podstatně varíují i přes mnohdy relativně vysoké počty participantů. Stále jsou však zřejmé výrazně zvýšené skóry klinických vzorků i podle pouhého rozpětí skóru v Tabulce 1 (1,9-15,5) a Tabulce 2 (bez pacientů v remisi: 16,3-30,8). V návaznosti na to Tabulka 2 ukazuje očividný rozdíl skóru mezi pacienty s trvajícím depresivním příznakem a pacienty v remisi, což je dobrá známka, ze které lze usuzovat na schopnost BDI-II rozlišovat nejen klinicky depresivní lidi od neklinické populace, ale rovněž i od dříve depresivních, kteří již tyto příznaky nevykazují.

Tabulka 2 – studie reportující skóry BDI-II klinických vzorků s unipolární depresí

Studie	Povaha vzorku	Věk	N	M	SD
Bizik et al. (2011)	hospitalizovaní	38,4±8,2	40	18,7	9,06
Bob et al. (2005)	hospitalizovaní s uni- či bipolární depresí	48,3±12,2	70	24,7	12,60
Bob et al. (2009)	ambulantní	35±10,2	40	20,7	11,99
Bob et al. (2010)	ambulantní	35±9,8	113	20,6	12,40
Bob et al. (2011)	ambulantní	36,3±9	40	16,3	10,50
Ocisková et al. (2014)	hospitalizovaní s lehkou až střední depresí	---	17	31,4	11,12
Ocisková et al. (2017)	hospitalizovaní/ambulantní	45,6±12,2	177	30,8	10,30
Preiss et al. (2009)	ambulantní v remisi	46,3±12,2	97	11,8	6,95
Preiss et al. (2010a)	ambulantní v remisi	46,8±12,7	31	7,4	6,18
Preiss et al. (2010b)	ambulantní v remisi	45,4±11,9	92	11,6	6,98
Rek (2017)	hospitalizovaní/ambulantní	49,1±11,5	44	30,2	11,80

Pro korektní diskuzi výsledků metaanalýzy by měly všechny studie pracovat s nezávislými vzorky, tzn. žádný participant by se neměl účastnit více než jedné studie. Přestože v žádné studii autoři explicitně neuvádí, že by využívali data získaná v rámci jiné z uvedených studií, lze si povšimnout některých velmi podobných deskriptivních charakteristik studií těchto autorů, například Ocisková et al. (2016; 2017) v Tabulce 1, či Preiss et al. (2009; 2010b) v obou tabulkách. Autory, jejichž studie se nacházejí v tabulkách vícekrát a obsahují podobné deskriptivní statistiky, jsem proto oslovil prostřednictvím e-mailu s prosbou o informace ohledně závislosti či nezávislosti výzkumných vzorků. Kontaktní autoři studií připustili možný překryv vzorků ve studiích Boba et al. (2009; 2010; 2011) a Preisse et al. (2009; 2010a; 2010b). Po komunikaci s autory a kontrole dostupných datasetů jsem se rozhodl z důvodu překryvu vzorku s jinými studiemi ze statistických metaanalýz vyřadit dvě studie Boba et al. (2009; 2011) a dvě studie Preisse et al. (2010a; 2010b). Údaje o vzorcích pacientů v remisi jsem do metaanalýz rovněž nezahrnul, jelikož tyto vzorky dle mého názoru nelze považovat za typickou součást populace klinické ani neklinické.

Metaanalytický design

Metaanalýzu jsem uskutečnil kvůli zřejmě odlišné podstatě populace na neklinických a klinických vzorcích zvlášť. Navíc jsem se po studiu Tabulky 1 rozhodl provést metaanalýzu neklinických vzorků zvlášť na studiích, kde byla metoda BDI-II administrována tváří v tvář, a tam, kde byla zadávána v elektronické online formě. Tato

volba dává smysl z mnoha důvodů, předně proto, že většina studií získávala své vzorky příležitostným výběrem a populace lidí ochotná se zúčastnit výzkumu online může být jiná než populace lidí ochotná nechat se vyšetřit výzkumníkem tvář v tvář. Za druhé je pak možné, že odlišná forma administrace ovlivňuje výsledné skóry – participant mohou při vyplňování elektronické online formy volit jiné strategie odpovídání. Rovněž místo a kontext, ve kterém participant odpovědi poskytuje, se od formy tvář v tvář liší. Třetím důvodem je interpretace výsledků analýzy, jelikož pro odhad populačních parametrů a cut-off skóre BDI-II v kontextu klinického vyšetření, které by vždy mělo probíhat individuálně tvář v tvář, je vhodné použít informace vzorků měřených v podobném kontextu.

Celkově jsem proto vypracoval tři dílčí metaanalýzy podle charakteru vzorku, od něhož údaje pochází: neklinický vzorek s BDI-II administrovanou tvář v tvář, neklinický vzorek s BDI-II administrovanou online a klinický vzorek (kterému bylo ve všech studiích BDI-II administrováno tvář v tvář).

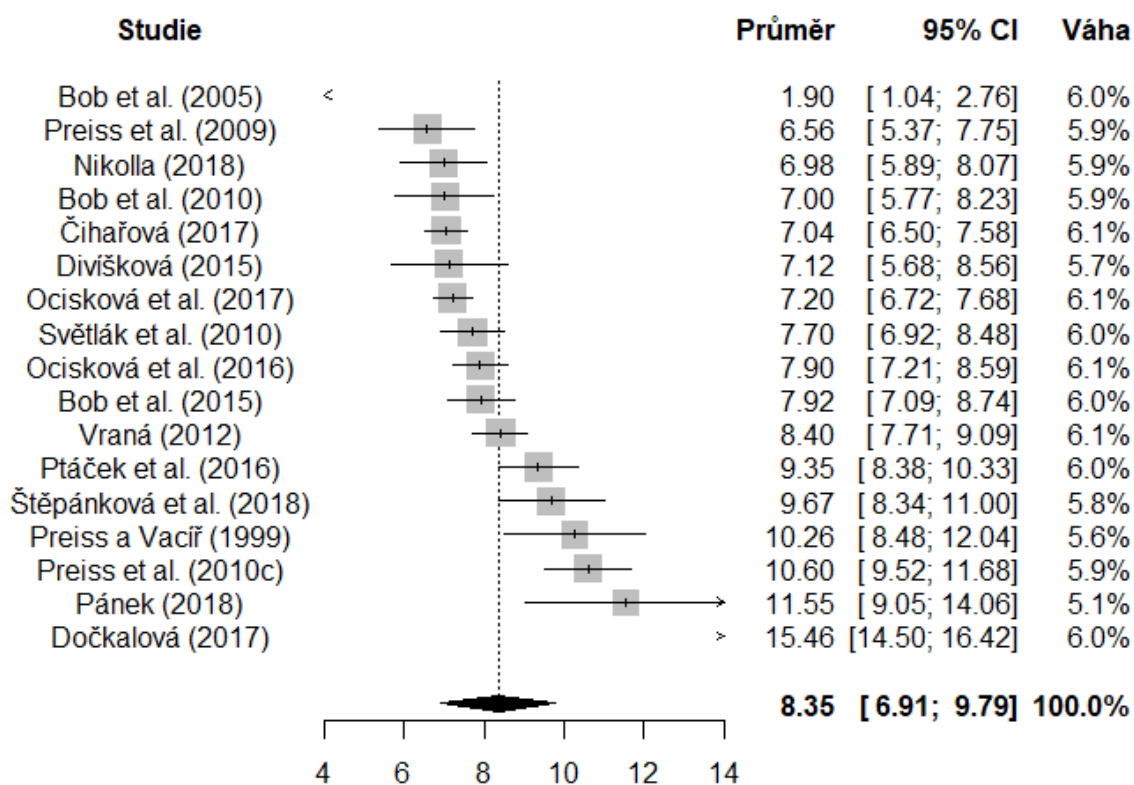
Pro všechny metaanalytické analýzy jsem zvolil model s náhodnými efekty. V kontextu této studie model s náhodnými efekty na rozdíl od modelu s fixními efekty předpokládá, že neexistuje jednotný průměrný skór BDI-II v populaci, nýbrž počítá s tím, že napříč populací mají hodnoty pravého skóru průměru přibližně normální rozložení. Průměrný populační skór se může lišit například v závislosti na pohlaví či vzdělání, jak bylo rozebráno v teoretickém úvodu. Model s náhodnými efekty navíc nepředpokládá, že by zkoumané studie byly přesnými replikacemi, ani že by tím pádem všechny studie měly reportovat přibližně stejné hodnoty (Cuijpers, 2016). Model s náhodnými efekty jsem se rozhodl použít, jelikož míru depresivity v analyzovaných studiích je možné podle mého mínění lépe charakterizovat jako rozložení průměrných skóre spíše než jako jedinou hodnotu průměru, jelikož by se mělo jednat o kontinuum depresivity v populaci v závislosti na rozdílných demografických charakteristikách a kontextových efektech administrace. Navíc je tento model schopen pracovat s variabilitou skóre mezi jednotlivými studiemi způsobenou například formou administrace, metodou výběru vzorku či specifickým prostředím, ve kterém byla studie provedena. To považuji pro účely této metaanalýzy, která sdružuje poznatky z nejrůznějších typů výzkumů, od diplomových prací, přes studie využívající BDI-II pouze okrajově, až po ty standardizační a psychometrické, za zjevnou výhodu. Intervaly spolehlivosti byly odhadnuty na základě metody Hartung-Knapp-Sidik-Jonkman, jelikož v simulačních studiích odhadovala přesněji oproti běžně používané metodě DerSimoniana a Lairda (IntHout, Ioannidis, & Borm, 2014). Pro odhad parametru τ^2 vyjadřující míru heterogenity mezi jednotlivými studiemi jsem zvolil metodu Restricted Maximum-Likelihood, která se zdá být i při zešíkmeném rozložení proměnné a malém počtu studií v porovnání s jinými metodami nejméně zkreslená a může být považována za zlatý standard (Langan et al., 2019).

Výsledky

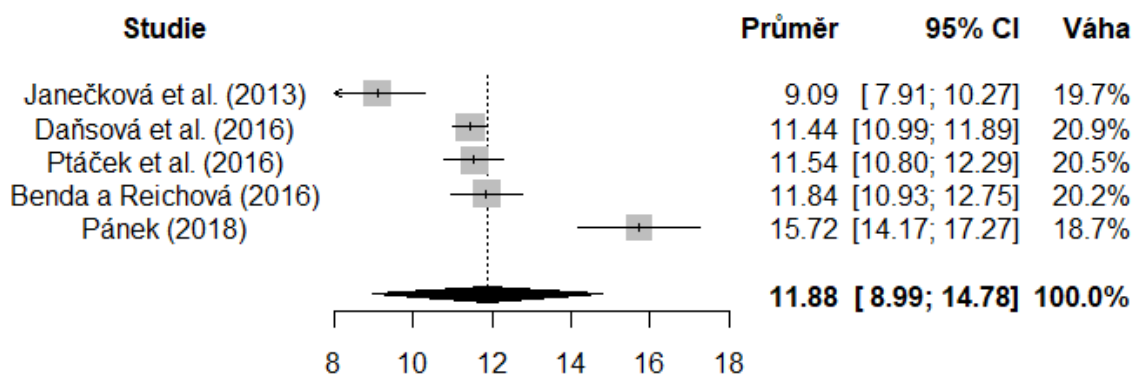
Veškeré statistické analýzy byly provedeny v programu R verze 3.5.2 (R Core Team, 2018). Použitá data jsou spolu se skriptem analýzy dostupná online na adrese <https://osf.io/9etj6>.

Jako první byla provedena metaanalýza dat neklinických vzorků, které BDI-II vyplňovaly metodou tužka-papír tváří v tvář s administrátorem. Tento první model zachycuje Obrázek 1. Na základě dostupných studií ($k = 17$) model odhaduje průměrný skór BDI-II při administraci tváří v tvář jako 8,35; 95% CI [6,91; 9,79], v modelu byly však i přes vyloučení studií s online formou administrace detekovány výrazně heterogenní vzorky ($\tau^2 = 7,55$; $Q(16) = 508,73$; $p < 0,001$).

Obrázek 1 – metaanalýza neklinických vzorků testovaných formou tváří v tvář



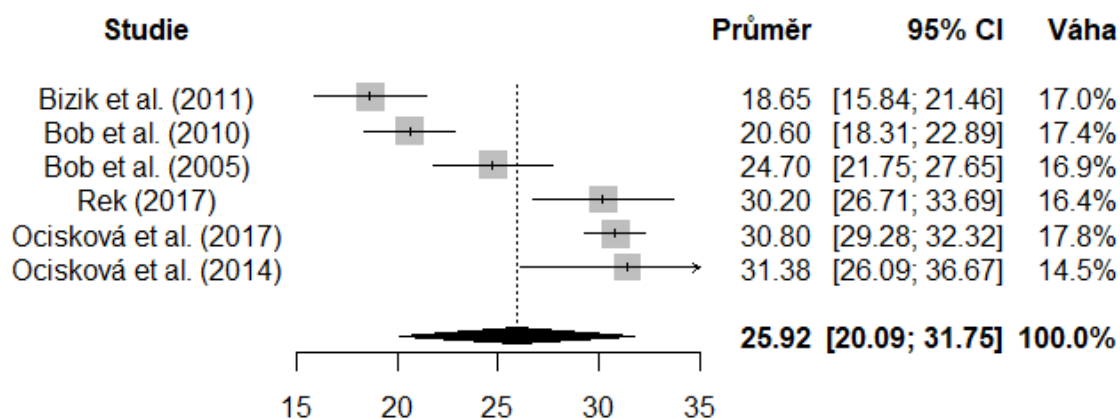
Obrázek 2 – metaanalýza neklinických vzorků testovaných online formou



Obrázek 2 představuje druhý model se studiemi, které administrovaly BDI-II online formou neklinickým vzorkům ($k = 5$). Tento model vykazuje o něco menší, avšak stále značnou heterogenitu vzorků ($\tau^2 = 4,91$; $Q(4) = 45$; $p < 0,001$). Odhadovaný populační průměr skóre BDI-II při administraci online činí 11,88; 95% CI [8,99; 14,78] a při porovnání obou modelů byl nalezen statisticky významný rozdíl v odhadovaných skórech ($Q(1) = 8,05$; $p = 0,005$). I přes malé množství studií ve druhém modelu a velkou míru heterogenity v obou z nich výsledky podporují domněnku, že forma administrace BDI-II ovlivňuje konečné skóre, a tudíž jejich rozdělení a analýza každé skupiny zvlášť má svoje opodstatnění.

Třetí model, který byl aplikován na klinické vzorky ($k = 6$), ukazuje opět velmi vysokou míru heterogenity ($\tau^2 = 28,35$; $Q(5) = 95,51$; $p < 0,001$). Model odhaduje průměrný skóre napříč klinickými vzorky na 25,92; 95% CI [20,09; 31,75], což je výrazně vyšší skóre než v neklinických vzorcích. Mezi neklinickými a klinickými vzorky je z hlediska odhadovaného průměru velmi významný rozdíl ($Q(1) = 50,36$; $p < 0,001$). Tento výsledek podporuje validitu BDI-II ve smyslu schopnosti odlišení zdravých a klinicky depresivních lidí. V Tabulce 3 pro přehlednost uvádím odhadované parametry skóre BDI-II pro různé populace. Směrodatná odchylka byla odhadnuta metodou průměru směrodatných odchylek jednotlivých studií vážených podle počtu participantů ve vzorku.

Obrázek 3 – metaanalýza klinických vzorků

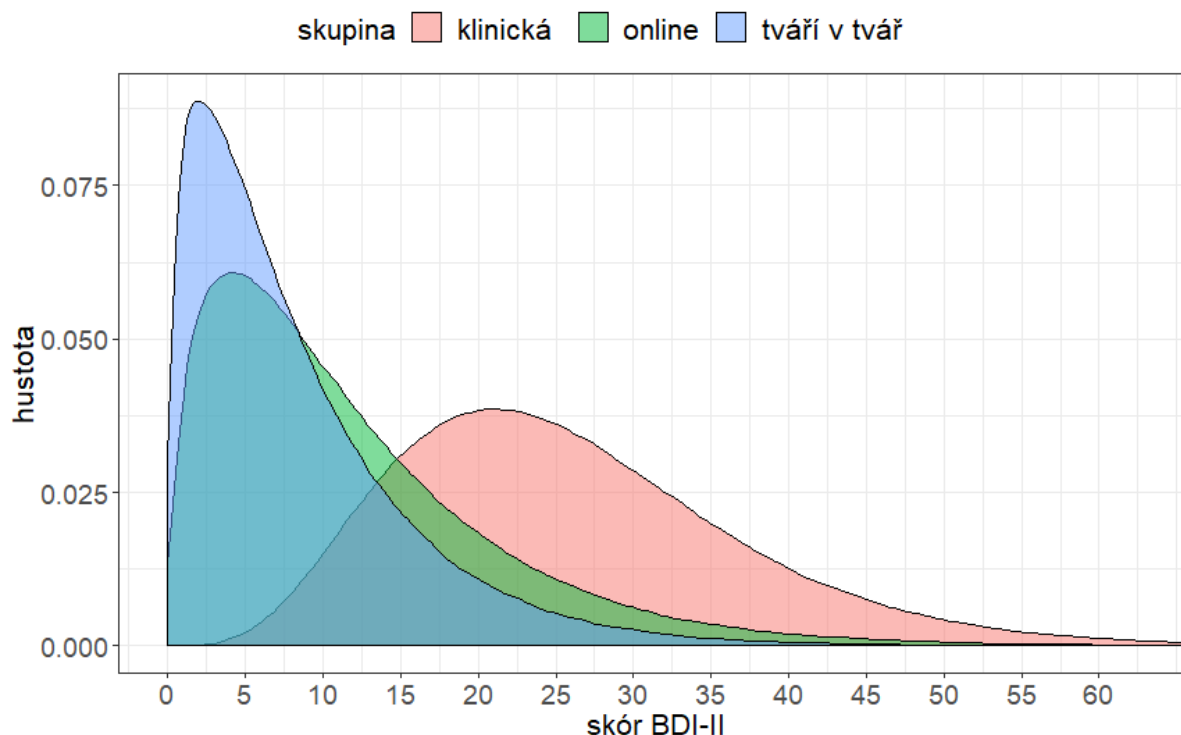


Tabulka 3 – přehled odhadovaných populačních parametrů na základě dostupných studií

Populace	M	95% CI	SD
Neklinická (tváří v tvář)	8,35	[6,91; 9,79]	7,3
Neklinická (online)	11,88	[8,99; 14,78]	9,62
Klinická	25,92	[20,09; 31,75]	11,23

Za účelem odhadu podoby populačního rozložení skóreů BDI-II jsem se rozhodl využít jako základ hodnoty průměrů a směrodatných odchylek získané z metaanalýzy. Dalším krokem bylo určit podobu rozložení. Po analýze dostupných primárních i sekundárních dat jsem dospěl k názoru, že minimálně u neklinické populace rozložení skóreů přibližně odpovídá podobě rozdělení chí-kvadrát. Charakteristickým znakem hrubých skóreů bylo výrazné pozitivní zešíkmení spolu s efektem podlahy vzhledem k povaze škály BDI-II, na které nelze získat nižší skóre než nula. Usoudil jsem proto, že rozložení chí-kvadrát poměrně přesně vystihuje deskriptivní statistiky výzkumných dat z neklinických vzorků, přičemž jde o matematicky přístupnou metodu simulace. Primární data z klinických vzorků se mi nepodařilo získat, simulace tak stojí na předpokladu, že skóre klinické populace víceméně odpovídají rozložení chí-kvadrát. V dalším kroku jsem proto nasimuloval pravděpodobnostní rozložení skóreů populací vždy tak, aby měla charakter rozdělení chí-kvadrát a jejich průměry a směrodatné odchylky se co nejvíce blížily hodnotám v Tabulce 3. Hustoty rozložení skóreů jsou znázorněny na Obrázku 4, ze kterého lze předběžně usuzovat na poměrně dobrou schopnost metody BDI-II rozlišovat mezi klinickou a neklinickou populací. Také lze vidět menší míru zešíkmení a vyšší průměrný skóre online formy administrace oproti formě tváří v tvář.

Obrázek 4 – graf hustoty rozložení simulovaných skóreů BDI-II



Tabulka 4 – cut-off skóry BDI-II na simulovaných datech

Cut-off skór	Senzitivita	1 - Specificita	Cut-off skór	Senzitivita	1 - Specificita
6/7	0,99	0,46	20/21	0,63	0,06
7/8	0,98	0,40	21/22	0,59	0,06
8/9	0,97	0,35	22/23	0,55	0,05
9/10	0,96	0,31	23/24	0,51	0,04
10/11	0,94	0,27	24/25	0,47	0,04
11/12	0,92	0,24	25/26	0,44	0,03
12/13*	0,90	0,20	26/27	0,41	0,03
13/14	0,87	0,18	27/28	0,37	0,02
14/15	0,84	0,16	28/29	0,34	0,02
15/16	0,81	0,13	29/30	0,31	0,02
16/17	0,78	0,12	30/31	0,29	0,01
17/18	0,74	0,10	31/32	0,26	0,01
18/19	0,71	0,09	32/33	0,24	0,01
19/20	0,67	0,07	33/34	0,21	<0,01

* optimální cut-off skór podle Youdenova indexu

Pro praktické klinické účely jsem na základě simulovaných dat stanovil cut-off skóry užitečné k rozlišení mezi klinicky depresivní a obecnou neklinickou populací, viz Tabulka 4. Pro analýzu jsem mimo simulovaných dat klinické skupiny použil pouze simulovaná data neklinické skupiny s formou administrace tváří v tvář. Důvodem byl fakt, že v klinickém kontextu se BDI-II administruje pouze tváří v tvář, a zahrnovat do analýzy data z populace testované online tak postrádá smysl. Tabulka pro každý cut-off skór uvádí očekávaný podíl správně diagnostikovaných klinicky depresivních a nesprávně diagnostikovaných klinicky nedepresivních lidí. Nejlepších výsledků podle Youdenova indexu dosahuje cut-off skór 12/13. Plocha pod křivkou (AUC) činí 0,92, což lze interpretovat jako argument ve prospěch velmi dobré rozlišovací schopnosti testu.

Diskuze

Již počáteční rešerše literatury upozornila na výrazné rozdíly mezi průměrnými skóry BDI-II napříč studiemi, což následné metaanalýzy potvrdily. Tyto rozdíly nelze v plném rozsahu uspokojivě vysvětlit vlivem pohlaví, věku ani vzdělání, jelikož první dvě charakteristiky nemají potenciál způsobit tak velké rozdíly a vzdělání vzorku se nezdá být systematicky jiné mezi studiemi s vysokými průměrnými skóry a nízkými. Významnou heterogenitu skóre v rozsahu až deseti bodů mezi klinickými vzorky lze podle mého názoru alespoň zčásti vysvětlit rozdílnou dobou léčby. Psychiatrickí pacienti, kteří jsou na konci léčby, by měli mít přibližně o pět bodů nižší skóry než pacienti na jejím začátku, jak ukazují výsledky Ociskové et al. (2014). V případě neklinických vzorků je však i při kontrole vlivu formy administrace příčina heterogenity nejasná. Usuzuji na to, že vnitřní rozdíly neklinických vzorků jsou důsledkem vlivu výběrového zkreslení vzorku a specifického kontextu administrace. Z Obrázku 1 a 2 odhaduji, že intervenující faktory mohou průměrné skóry vychýlit o tři body i více. Pro výzkumné i praktické uživatele

BDI-II proto doporučuji provádět administraci na reprezentativním vzorku v konzistentní formě a zajistit co možná nejrovnější podmínky pro všechny vyšetřované, jelikož kontextuální vlivy mají pravděpodobně potenciál výrazně zkreslit výsledné skóry. Je otázkou pro budoucí výzkum, do jaké míry hrají tyto faktory roli při zkreslování výsledků i jiných psychodiagnostických metod a do jaké míry mohou být výsledky standardizačních i běžných studií jejich vlivem zkreslené.

Výsledky porovnání neklinických skupin mezi sebou ukazují, že forma administrace ovlivňuje skór BDI-II, přičemž skupina s online formou dosahuje průměrného skóru o přibližně tři a půl bodu výše než skupina s administrací tvář v tvář. To je v těchto podmínkách ekvivalentem Cohenova $d = 0,41$, což ukazuje na potenciální slabou až střední sílu účinku, nicméně je třeba brát v potaz zkreslení Cohenova d vyplývající ze zešíkmeného rozložení skórů. Domnívám se, že v rámci praktické interpretace je proto srozumitelnější zůstat v diskuzi na úrovni hrubých skórů. V každém případě odhadovaný průměrný rozdíl více než třech bodů představuje nezanedbatelný faktor zvláště pro studie, které kombinují data získaná online a tvář v tvář, jelikož systematický nárůst skórů části vzorku může zakrýt skutečné vztahy proměnných či naopak může vztahy uměle vytvořit. Na to je potřeba brát ohled i v rámci ústní a písemné administrace BDI-II, jak uvádějí Štěpánková Georgi et al. (2018), kteří našli na vzorku seniorů o více než šest bodů nižší průměrné skóry ústní formy administrace.

Pro rozlišení klinicky depresivní a neklinické populace v rámci praktické psychodiagnostiky Youdenův index označil za nejefektivnější cut-off skór 12/13 s 90% senzitivitou a 80% specificitou, což jsou vzhledem k povaze BDI-II dle mého názoru velmi uspokojivé hodnoty. Tento cut-off skór je nižší než hodnota 16/17 podle Ociskové et al. (2017), což je nejspíše zapříčiněné poměrně vysokými BDI-II skóry pacientů a stejně tak poměrně nízkými skóry kontrolní skupiny v jejich studii oproti jiným vzorkům, které byly v méj metaanalýze a potažmo pak v simulaci skórů rovněž zahrnuty, viz Obrázek 1 a 3. Ve shodě s Ociskovou et al. (tamtéž) doporučuji cut-off skór 12/13 brát jako výchozí bod pro screening zhoršeného duševního stavu, nikoliv deprese. Zároveň pro praktickou psychodiagnostiku doporučuji v závislosti na hodnotách v Tabulce 4 stanovovat vlastní cut-off skóry podle účelu a možností vyšetření. Chce-li vyšetřující zjišťovat přítomnost mírnějších depresivních symptomů, lze nastavit cut-off skór 12/13, případně i nižší. Pro detekci výrazně depresivních lidí je žádoucí posunout hranici naopak výše a zajistit tak lepší specificitu. Navíc, jak jeden z recenzentů správně poukázal, použití cut-off skóru na základě Youdenova indexu funguje optimálně při 50% prevalenci nemoci. Jestliže tedy pracujeme se vzorkem, ve kterém je odhadovaná prevalence nižší, je záhodno stanovit vyšší cut-off skór, abychom předešli zvýšené míře falešně pozitivních diagnóz, a vice versa.

Navrhovaný ideální cut-off skór je nicméně podobný cut-off skórům navrhnutým v zahraničních metaanalýzách. Erford, Johnson a Bardoshi (2016) uvádí jako optimální hodnotu právě 13, jakožto medián cut-off skórů ze vzorku 20 diagnostických studií. Kombinací údajů z 26 studií pak na souhrnném vzorku téměř 14000 participantů autoři

(tamtéž) dospívají k odhadovanému průměrnému skóru neklinické populace 8,39 se směrodatnou odchylkou 6,87, což je velice blízko hodnotám, které jsem v rámci této studie odhadnul pro neklinickou populaci testovanou tvář v tvář ($M = 8,35$; $SD = 7,3$). Wang a Gorenstein (2013) dodávají, že pro zjišťování deprese v neklinickém vzorku jsou obecně nejúčinnější cut-off skóry v rozsahu 10-16, v klinickém vzorku 7-20 a ve vzorku hospitalizovaných pacientů s psychiatrickou diagnózou 19-31. I toto dokládá důležitost rozvahy nad volbou vhodného cut-off skóru v závislosti na povaze vzorku populace, pro který diagnostiku provádíme.

Závěrem chci připomenout, že BDI-II je určena pro diagnostiku závažnosti depresivních symptomů, nikoliv jako nástroj pro určení klinické diagnózy deprese, a že výše vypsané cut-off skóry mají sloužit k tomuto účelu jako obecná a srozumitelná vodítka, v žádném případě jako tvrdé hranice. V neposlední řadě je při praktické psychodiagnostice pomocí BDI-II nezbytné přihlížet k obsahu jednotlivých položek. Vysoký skór v položce „Sebevražedné myšlenky nebo přání“ může mít odlišné příčiny, nemluvě o závažnosti praktických dopadů na život diagnostikovaného, než vysoký skór v položce „Nerozhodnost“, byť oba skóry mají na součtové škále stejnou váhu. Je dobré pamatovat i na to, že vysoký skór položky nutně neznamena vyšší depresivitu, například vysoký skór v položce „Změny v chuti k jídlu“ může mít zdroj v nesouvisejícím fyziologickém onemocnění, probíhající dietě, apod. Upřímně proto doporučuji při podezření na zvýšenou depresivitu využít krátkosti a jednoduchosti BDI-II a s vyšetřovaným alespoň stručně diskutovat jeho odpovědi, čímž lze získat cenné diagnostické informace.

Limity studie

Všechny výsledky metaanalýz mohou být zkreslené kvůli předpokladu normálního rozložení, které je u BDI-II skórů pozitivně zešikmené. Rovněž je třeba považovat výsledky pouze za odhady, jelikož jsou založeny na poměrně malém a navíc výrazně heterogenním vzorku studií a analýze sekundárních, nikoliv primárních, dat. Kvůli nepřístupnosti velké části primárních dat klinických vzorků je možné, že simulační skóry klinické populace nevystihují věrně rozložení skutečných skórů, a odhadované cut-off skóry tak mohou být do neznámé míry zkreslené.

Na základě připomínek recenzentů zde považuji za nutné zmínit i fakt, že analýzy mohly být zkresleny větším podílem validizačních či jiných studií oproti normativním. Zkreslení a vysoká heterogenita může pramenit z potenciálních rozdílů mezi vzorky kvůli odlišné povaze studií. K tomu může přispívat často nejasný charakter kontrolního souboru, ze kterého byli v některých studiích lidé s depresivními symptomy vyřazováni a v jiných ne. Také zkreslení vlivem základních demografických proměnných je možné, nicméně vzhledem k nedostupnosti velké části primárních dat není možné velikost tohoto zkreslení přesně odhadnout, vyjma domněnek zmíněných v teoretickém úvodu.

Poděkování

Děkuji všem autorům, kteří se mnou laskavě sdíleli informace i data ze svých studií. Rovněž děkuji anonymním recenzentům a doc. Stanislavu Ježkovi za podnětnou zpětnou vazbu k manuskriptu.

Zdroje

- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. San Antonio, 78(2), 490–498.
- Benda, J., & Reichová, A. (2016). Psychometrické charakteristiky české verze Self-Compassion Scale (SCS-CZ). *Československá psychologie*, 60(2), 120–136.
- Bizik, G., Bob, P., Raboch, J., Světlák, M., Šimek, J., Pec, O., ... Zima, T. (2011). Dissociation and Immune Dysregulation: A Preliminary Report. *Activitas Nervosa Superior*, 53(3–4), 141–145. <https://doi.org/10.1007/BF03379937>
- Bob, P., Jasová, D., & Raboch, J. (2011). Subclinical Epileptiform Process in Patients with Unipolar Depression and Its Indirect Psychophysiological Manifestations. *PLoS ONE*, 6(11), e28041. <https://doi.org/10.1371/journal.pone.0028041>
- Bob, P., Sélešová, P., Raboch, J., & Kukla, L. (2015). Dissociative Symptoms and Mother's Marital Status in Young Adult Population: *Medicine*, 94(2), e408. <https://doi.org/10.1097/MD.0000000000000408>
- Bob, P., Šusta, M., Gregusová, A., & Jasová, D. (2009). Dissociation, cognitive conflict and nonlinear patterns of heart rate dynamics in patients with unipolar depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 33(1), 141–145. <https://doi.org/10.1016/j.pnpbp.2008.11.005>
- Bob, P., Šusta, M., Gregusová, A., Jasová, D., Raboch, J., & Mishara, A. (2010). Traumatic Stress, Dissociation, and Limbic Irritability in Patients with Unipolar Depression Being Treated with SSRIs. *Psychological Reports*, 107(3), 685–696. <https://doi.org/10.2466/02.15.16.PR0.107.6.685-696>
- Bob, P., Šusta, M., Pavlát, J., Hynek, K., & Raboch, J. (2005). Depression, traumatic dissociation and epileptic-like phenomena. *Neuroendocrinology Letters*, 26(4), 321–326.
- Cuijpers, P. (2016). Meta-analyses in mental health research: A practical guide. Získáno z <https://indd.adobe.com/view/5fc8f9a0-bf1e-49d3-bf5f-a40bfe5409e0>
- Čihařová, M. (2017). Deprese a Beckova škála deprese: Normativní studie české verze a srovnání se zahraničními daty (Diplomová práce, Univerzita Karlova, Filozofická fakulta, Katedra psychologie, Praha). Získáno z <https://dspace.cuni.cz/handle/20.500.11956/86567>
- Daňsová, P., Masopustová, Z., Hanáčková, V., Kicková, K., & Korábová, I. (2016). Metoda Patient Health Questionnaire-9: Česká verze. *Československá psychologie*, 60(5), 468–481.
- Divíšková, P. (2015). Vztah místa kontroly k vnímání času a depresivitě u klientů léčících se s diagnózou F10 (Univerzita Palackého v Olomouci, Filozofická fakulta, Olomouc). Získáno z <https://theses.cz/id/z1w8y6/>

- Dočkalová, M. (2017). Interpretace Testu hvězd a vln v kontextu vybraných testů osobnosti (Diplomová práce, Masarykova univerzita, Filozofická fakulta, Brno). Získáno z <https://is.muni.cz/th/i9pj8/>
- Erford, B. T., Johnson, E., & Bardoshi, G. (2016). Meta-Analysis of the English Version of the Beck Depression Inventory–Second Edition. Measurement and Evaluation in Counseling and Development, 49(1), 3–33. <https://doi.org/10.1177/0748175615596783>
- Gottfried, J. (2015). Beckova sebesposuzovací škála depresivity pro dospělé: Recenze metody. Testforum, 4(5). <https://doi.org/10.5817/TF2015-5-35>
- IntHout, J., Ioannidis, J. P., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Medical Research Methodology, 14(1), 25. <https://doi.org/10.1186/1471-2288-14-25>
- Janečková, D., Dostál, D., & Plháková, A. (2013). Kvalita spánku, cirkadiánní preference a zdravý životní styl u vysokoškolských studentů. Praktický lékař, 93(3), 114–120.
- Langan, D., Higgins, J. P. T., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., ... Simmonds, M. (2019). A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. Research Synthesis Methods, 10(1), 83–98. <https://doi.org/10.1002/jrsm.1316>
- Nikolla, M. (2018). Příspěvek k validaci české verze Starksteinovy škály apatie (Diplomová práce, Pražská vysoká škola psychosociálních studií, Praha). Získáno z https://www.pvpsps.cz/data/2018/11/07/12/nikolla_martina.pdf
- Ocisková, M., Praško, J., Kamarádová, D., Grambal, A., Látalová, K., & Sigmundova, Z. (2014). Relationship Between Internalized Stigma and Treatment Efficacy in Mixed Neurotic Spectrum and Depressive Disorders. Neuroendocrinology Letters, 35(8), 101–107.
- Ocisková, M., Praško, J., Kupka, M., Marackova, M., Latalova, K., Činčulová, A., ... others. (2017). Psychometric evaluation of the Czech Beck Depression Inventory-II in a sample of depressed patients and healthy controls. Neuroendocrinology Letters, 38(2), 98–106.
- Ocisková, M., Sobotková, I., Praško, J., & Mihal, V. (2016). Measuring hope: Standardization of the Czech version of the Adult Dispositional Hope Scale in healthy adults. Neuroendocrinology Letters, 37(8), 543–550.
- Pánek, P. (2018). Psychometrická studie Beckovy škály deprese (Diplomová práce, Masarykova univerzita, Filozofická fakulta, Brno). Získáno z <https://theses.cz/id/ifge6t/>
- Preiss, M., Kramská, L., Dočkalová, E., Holubová, M., & Kučerová, H. (2010). Attentional networks in euthymic patients with unipolar depression. European Psychiatry, 25(2), 69–74. <https://doi.org/10.1016/j.eurpsy.2009.08.007>
- Preiss, M., Kučerová, H., Lukavský, J., Sos, P., Štěpánková, H., & Čermáková, R. (2010). Cognitive deficits in hospitalized and never hospitalized remitted unipolar depressive patients. The European Journal of Psychiatry, 24(3), 129–135.
- Preiss, M., Kučerová, H., Lukavský, J., Štěpánková, H., Sos, P., & Kawaciukova, R. (2009). Cognitive deficits in the euthymic phase of unipolar depression. Psychiatry Research, 169(3), 235–239. <https://doi.org/10.1016/j.psychres.2008.06.042>
- Preiss, M., & Vacíř, K. (1999). Beckova sebesposuzovací škála depresivity pro dospělé: BDI-II. Příručka. Brno: Psychodiagnostika.

- Preiss, M., Vraná, K., & Kuchařová, J. (2010). Depresivita a úzkost v průběhu života v návaznosti na pojetí Václava Příhody. *Československá psychologie*, 54(4), 357–365.
- Ptáček, R., Raboch, J., Vňuková, M., Hlinka, J., & Anders, M. (2016). Beckova škála deprese BDI-II-standardizace a využití v praxi. *Česká a Slovenská Psychiatrie*, 112(6), 270–274.
- R Core Team. (2018). R: A language and environment for statistical computing (Verze 3.5.2). Získáno z <https://www.R-project.org/>
- Rek, O. (2017). Vztah mezi sebestigmatizací a strategiemi zvládání stresu u lidí s depresivní poruchou (Diplomová práce, Univerzita Palackého v Olomouci, Filozofická fakulta, Olomouc). Získáno z <https://theses.cz/id/zduhyp/>
- Světlák, M., Bob, P., & Kukleta, M. (2010). Complex partial seizure-like symptoms and smoking in university students. *Scripta Medica Facultatis Medicae Universitatis Brunensis Masarykianae*, 83, 124–129.
- Štěpánková Georgi, H., Horáková Vlčková, K., Lukavský, J., Kopeček, M., & Bareš, M. (2018). Beck Depression Inventory-II: Self-report or interview-based administrations show different results in older persons. *International Psychogeriatrics*, 1–8. <https://doi.org/10.1017/S1041610218001187>
- Urbánek, T. (2010). Nejpoužívanější psychodiagnostické metody v České republice. *Testforum*, 1(1), 6–9. <https://doi.org/10.5817/TF2010-1-3>
- Vraná, K. (2012). Subjektivně vnímaná depresivní symptomatika u dospělé české populace—Využití Beckova inventáře deprese (Diplomová práce, Univerzita Karlova, Filozofická fakulta, Katedra psychologie, Praha). Získáno z <https://dspace.cuni.cz/handle/20.500.11956/44783>
- Wang, Y.-P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria*, 35(4), 416–431. <https://doi.org/10.1590/1516-4446-2012-1048>
- Whisman, M. A., Judd, C. M., Whiteford, N. T., & Gelhorn, H. L. (2013). Measurement Invariance of the Beck Depression Inventory–Second Edition (BDI-II) Across Gender, Race, and Ethnicity in College Students. *Assessment*, 20(4), 419–428. <https://doi.org/10.1177/1073191112460273>

Jaroslav Gottfried (2019): Czech normative scores of Beck Depression Inventory: a meta-analysis

Abstract: *The article introduces the history of local adaptation and standardization of Beck Depression Inventory together with sociodemographic variables that have a potential to affect its scores. The core of the article consists of the overview of existing studies in which Czech adaptation of the method has been administered to healthy or clinically depressed participants. Resulting descriptives were used in a meta-analysis in order to estimate the properties of population scores. The meta-analysis suggests that the form of administration affects sum scores. Specifically, administration via online questionnaires raises sum scores by more than three points, on average. Even though the form of administration has been controlled for, mean scores and standard deviations among samples were strongly heterogeneous, which complicates the estimation of population scores. Sampling bias and administration context effects are the probable causes of these differences. Due to the strong heterogeneity of mean scores among studies, it is advisable to check for these potential biases in other clinical methods, too. Based on theoretical assumptions and the estimates of population scores, the distribution of population scores has been simulated. This distribution has been used to create cut-off scores for practical use in psychodiagnostics. Lastly, the meta-analytic results are compared with those of international meta-analyses and recommendations for employing Beck Depression Inventory in research and practice are made. The implications of this study are hindered by a low number of studies analyzed, the heterogeneity of scores among studies, and the distribution of scores not being normal.*

Keywords: *Beck Depression Inventory; BDI-II; depression; meta-analysis; norms*